

Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (*DGAT1*) with variation at a quantitative trait locus for milk fat content

Andreas Winter*, Wolfgang Krämer*, Fabian A. O. Werner*, Sonja Kollers*, Srinivas Kata†, Gregor Durstewitz*, Johannes Buitkamp‡, James E. Womack†, Georg Thaller*, and Ruedi Fries*[§]

*Lehrstuhl für Tierzucht der Technischen Universität München, 85350 Freising-Weißenstephan, Germany; †Department of Veterinary Pathobiology, Texas A&M University, College Station, TX 77843-4463; and ‡Bayerische Landesanstalt für Tierzucht, Grub, 85586 Poing, Germany

Contributed by James E. Womack, May 16, 2002

***DGAT1* encodes diacylglycerol *O*-acyltransferase (EC 2.3.1.20), a microsomal enzyme that catalyzes the final step of triglyceride synthesis. It became a functional candidate gene for lactation traits after studies indicated that mice lacking both copies of *DGAT1* are completely devoid of milk secretion, most likely because of deficient triglyceride synthesis in the mammary gland. Our mapping studies placed *DGAT1* close to the region of a quantitative trait locus (QTL) on bovine chromosome 14 for variation in fat content of milk. Sequencing of *DGAT1* from pooled DNA revealed significant frequency shifts at several variable positions between groups of animals with high and low breeding values for milk fat content in different breeds (Holstein-Friesian, Fleckvieh, and Braunvieh). Among the variants was a nonconservative substitution of lysine by alanine (K232A), with the lysine-encoding allele being associated with higher milk fat content. Haplotype analysis indicated the lysine variant to be ancestral. Two animals that were typed heterozygous (*Qq*) at the QTL based on marker-assisted QTL-genotyping were heterozygous for the K232A substitution, whereas 14 animals that are most likely *qq* at the QTL were homozygous for the alanine-encoding allele. An independent association study in Fleckvieh animals confirmed the positive effect of the lysine variant on milk fat content. We consider the nonconservative K232A substitution to be directly responsible for the QTL variation, although our genetic studies cannot provide formal proof.**

Milk fat content in cattle is a continuously distributed trait with heritability estimates between 0.45 and 0.50 (1). The systematic mapping of quantitative trait loci (QTL) underlying the genetic variance of milk production traits resulted in approximate map positions of QTL for milk fat content (2–5). The most consistent results were reported for a QTL on chromosome 14 (4, 6). On the basis of the analysis of seven chromosomes carrying the milk fat content increasing allele *Q*, the QTL was pinpointed to an interval of 5 centimorgans (cM) (7). Even with further refinement of the QTL position, determining the molecular basis of the variation at the QTL by a purely positional approach would be extremely difficult. The combined information from the function of gene products and the chromosomal location of genes allows the identification of functional and positional candidate genes which can then be screened for DNA variation that might represent the molecular basis for the variation at the QTL.

Recently, Cases *et al.* (8) showed at least two enzymes catalyze the reaction in which diacylglycerol is covalently joined to long-chain fatty acyl-CoAs to form triglycerides as major constituents of fat, including fat of secreted milk. *DGAT1* (diacylglycerol *O*-acyltransferase homolog 1, mouse) encodes one of these enzymes (EC 2.3.1.20). It became a functional candidate gene for milk fat content after studies indicated that lactation is

deficient in female mice lacking both copies of *DGAT1*, most likely because of impaired or deficient triglyceride synthesis in the mammary gland (9). In the present study we report the mapping of *DGAT1* in the region of the milk fat content QTL on bovine chromosome 14. We present evidence that increased milk fat content in different breeds is strongly associated with a lysine at position 232 of the protein encoded by bovine *DGAT1*; an alanine at this position is associated with lowered milk fat content.

Materials and Methods

Primers. Primers were designed based on expressed sequence tag (EST)-derived bovine DNA sequence, human cDNA sequence, and obtained bacterial artificial chromosome (BAC) sequence by using the PRIMER3 program (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.www.cgi>).

Screening of BAC Library. The bovine RPCI-42 genomic BAC library (10) was screened by using a radiolabeled PCR product of 565 bp containing parts of bovine exons 16 and 17 as a probe for colony hybridization. Three clones (56-F1, 240-A1, 269-H17) were identified and confirmed by PCR.

Mapping. The map position of *DGAT1* on human chromosome 8 was determined by PCR analysis of the Gene Bridge 4 human-hamster radiation hybrid panel (11) with primers derived from the 3' untranslated region of human *DGAT1* mRNA (F, 5'-GAGGCTCTCTGCCCTATG-3', R, 5'-TTTATTGACACCCTCGGACC-3'). Results were analyzed with a program from The Sanger Center (<http://www.sanger.ac.uk/Software/RHserver/RHserver.shtml>). The chromosomal location of *DGAT1* in cattle was determined by fluorescence *in situ* hybridization observing a standard protocol and by PCR analysis of a bovine-hamster radiation hybrid panel (12) with the primers: F, 5'-GTGGCTGTCACTCATCATCG-3' and R, 5'-TGTGCA-CAGCACTTTATTGAC-3'.

DNA Samples and Pooling. DNA was prepared from bovine semen by a standard method. Semen samples were obtained from individuals belonging to different taurine (*Bos taurus taurus*) and indicine (*B. taurus indicus*) breeds. DNA concentrations were measured by using a fluorometer and adjusted to a concentration

Abbreviations: DGAT, diacylglycerol acyltransferase; DGAT1, symbol of gene encoding DGAT; QTL, quantitative trait locus; BAC, bacterial artificial chromosome; K232A, lysine by alanine substitution at amino acid position 232 of DGAT; EST, expressed sequence tag; RFLP, restriction fragment length polymorphism.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AJ318490).

[§]To whom reprint requests should be addressed. E-mail: ruedi.fries@tierzucht.tum.de.

of 25 ng/ μ l. Quality and quantity of DNA samples were assessed independently through agarose gel electrophoresis and PCR. Only DNA samples allowing consistent quantification yielding equal PCR amplification were used for the preparation of DNA pools. Pools were composed of DNA from bulls of German Holstein-Friesian (32 animals per pool), German Fleckvieh (32 animals per pool) and German Braunvieh (20 animals per pool) populations with highest (+) and lowest (–) breeding values for milk fat content (percentage of fat in milk) as evaluated internationally by INTERBULL (<http://www.interbull.org>) by using the Best Linear Unbiased Prediction (13). Holstein-Friesian bulls were selected from 2,857 artificial insemination bulls born 1988 or later (mean breeding value for milk fat content of -0.148 ± 0.284 SD). Bulls with breeding values greater than 0.48 and less than -0.68 were selected. Mean breeding values of pooled groups were: HF32+, 0.622 ± 0.125 ; HF32–, -0.771 ± 0.063 . Fleckvieh bulls were selected among 4,070 artificial insemination bulls born 1990 or later (mean breeding value of 0.089 ± 0.217). Bulls with breeding values greater than 0.56 and less than -0.31 were selected. Mean breeding values of pooled groups were: FV32+, 0.683 ± 0.153 ; FV32–, -0.454 ± 0.061 . Braunvieh bulls were selected among 656 artificial insemination bulls born 1990 or later (mean breeding value of 0.006 ± 0.185). Bulls with breeding values greater than 0.29 and less than -0.22 were selected. Mean breeding values of pooled groups were: BV20+, 0.424 ± 0.156 ; BV20–, -0.317 ± 0.096 .

Polymorphism Detection. Screening for variation was performed by PCR amplification and sequencing on an ABI377 sequencer by using DNA samples of individual animals and pooled DNA and by using primer sets covering *DGAT1* and the flanking ends (for primer sequences please contact authors). Sequence data were analyzed by using the Phred/Phrap/Polyphred/Consed software suite (14–17). A GC-to-AA exchange at positions 10433–10434 resulted in preferential PCR amplification of the AA allele. The amplitudes of the sequencing traces varied considerably between heterozygous animals. The inclusion of 5% DMSO in the PCR resulted in equal amplification of the two alleles and consistent patterns for the heterozygous samples.

Estimation of Allele Frequencies on the Basis of Sequencing Traces. Amplitude values at variable positions were extracted from data files “poly” created by the base calling program Phred (14, 15). The amplitude value for a given base was divided by the normalization factor for that base. Comparisons between the normalized amplitude values of pooled DNA (P) and mean normalized amplitude values of homozygous (Ho) and heterozygous (He) animals, respectively, were used to estimate allele frequencies. First, peak drops of pooled DNA (DP) and of heterozygous individuals (DH) relative to the peaks of homozygous individuals were calculated, i.e., $DP = Ho - P$, $DH = Ho - He$. Next, the following quotients were calculated: $Q1 = P/Ho$, $Q2 = P/He$ for $P \leq He$, and $Q2 = He/P$ for $P > He$, $Q3 = DP/DH$ for $DP \leq DH$, and $Q3 = DH/DP$ for $DP > DH$. The largest quotient dictated the calculation method for estimating optimal allele frequency (F) when assuming a linear relationship between normalized amplitude values and frequency in the given data range. The calculation methods were: $F = P/Ho$ for Q1 being the largest quotient; $F = (P \times 0.5)/He$ for Q2 being the largest quotient; $F = 1 - [(DP \times 0.5)/DH]$ for Q3 being the largest quotient. The estimates for base 1 and the alternative base 2 were designated F1 and F2, respectively. Finally, an average estimate for each base was calculated: $F1_{average} = [(1 - F2) + F1]/2$ and $F2_{average} = 1 - F1_{average}$.

QTL Mapping. QTL mapping was performed in a Fleckvieh granddaughter design (18) consisting of 20 paternal half-sib families with a total of 695 sons (18–58 sons per sire). Obser-

vations used for linkage analysis were INTERBULL breeding values for milk fat content. Breeding values of sons had an average reliability of 0.85 and ranged from -0.60 to $+0.96$ (mean of 0.027 ± 0.251 SD). Sires and their sons were genotyped for eight microsatellite markers on chromosome 14. The *ILSTS039* upstream primer was newly designed (5'-CCGAACAGAGTTGGACACAA-3'). All other microsatellites were amplified by using published primer pairs (<http://www.ri.bbsrc.ac.uk>, <http://www.marc.usda.gov>) and analyzed by multiplexing fluorescence-labeled PCR products on an ABI310 DNA sequencer. Average marker heterozygosity of sires was 62.5% with a maximum of 90% at the position of the marker *ILSTS039*. A standard interval-mapping procedure with multiple linked markers (19) was applied for linkage analysis. A chromosome-wide significance threshold was obtained empirically based on 10,000 permutations of the breeding values as described (20). The genome-wide significance threshold accounting for multiple chromosomes was determined by Bonferroni correction (21).

Restriction Fragment Length Polymorphism (RFLP) Assay. Detection of allelic variation at nucleotide positions 10433–10434 of the *DGAT1* gene was performed by *CfrI* cleavage in a 411-bp PCR product (primers: F, 5'-GCACCATCTCTTCCCTCAAG-3'; R, 5'-GGAAGCGCTTTCGGATG-3'). Cleavage by *CfrI* is diagnostic for the alanine-bearing allele. Five percent DMSO was added to the PCR reaction for equal amplification of both alleles.

Assessment of a VNTR in the 5' End of *DGAT1*. The number of repetitions of the repeat unit was determined on an ABI377 sequencer on the basis of the electrophoretic mobility of fluorescence-labeled PCR products (primers: F, 5'-TCAGGATC-CAGAGGTACCAG-3'; R, 5'-GGGGTCCAAGGTTGATACAG-3').

Results

To assess the positional candidate gene status of *DGAT1*, its precise position on human chromosome 8 was determined. Radiation hybrid mapping placed it next to marker AFMa082wh9 at cR3000 = 552.63 (lod score = 9.5) on the GB4 map (<http://www.ncbi.nlm.nih.gov/genemap99/>), 3.87 centirads from *KIAA0278*, a marker that is 16.6 centirads from *CSSM066* on the bovine radiation hybrid map (7). Ref. 7 showed *CSSM066* to be part of the identical-by-descent segment that indicated the minimal mapping interval of the milk fat content QTL, which provided indirect support of the positional candidate status of *DGAT1* with regard to the QTL.

BLASTN analysis with human *DGAT1* mRNA sequence identified 12 bovine sequences in the EST division GenBank. Some of these ESTs have been annotated as *DGAT1*-specific by the COMPASS approach (22). The EST-derived bovine mRNA corresponds to the human mRNA from position 509 to the 3' end.

Bovine BAC clones carrying *DGAT1* were isolated to obtain the sequence of the 5' end of the gene and its promoter region and to determine its structure. Iterative sequencing directly from BAC-DNA, sequencing from genomic DNA, and comparison of the resulting sequence with human mRNA and bovine EST-derived mRNA for DGAT was the basis for unraveling the exon-intron structure of the gene (Fig. 1). In total, the determined sequence spans 14,117 bp. Mapping of BAC-DNA containing *DGAT1* by fluorescence *in situ* hybridization assigned bovine *DGAT1* to a region on bovine chromosome 14 close to the centromere, the approximate physical location of the milk fat content QTL. Radiation hybrid mapping directly in cattle placed *DGAT1* distal to *ILSTS039*, a marker that at the time indicated the proximal boundary of the chromosome segment predicted to contain the QTL (7). This location of *DGAT1* theoretically excluded it as a positional candidate for the milk fat content

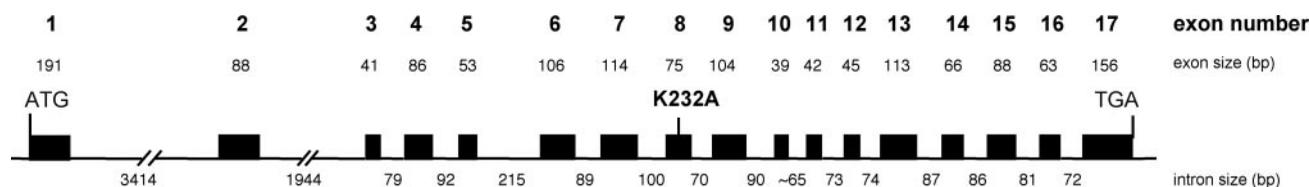


Fig. 1. Structure of bovine *DGAT1*. Exons are shown as rectangles. The sequence of intron 10 (≈ 65 bp) could not be completely resolved by sequencing. K232A indicates the only nonsynonymous mutation.

QTL. However, because of the potential errors in determining a QTL interval by identical-by-descent-mapping, we continued to treat *DGAT1* as a candidate gene.

We screened pooled DNA and DNA of individuals from different breeds and species for variation within the determined sequence. Direct sequencing revealed 19 variable positions and regions (Table 1), 17 of which were single-base exchanges, whereas two represented variable numbers of repeat units. We also observed variable PCR amplification in a region that could not be resolved by sequencing. However, fragmentary sequence information suggests the presence of a longer G stretch that could be variable in length. Variable positions 10433 and 10434 are located in exon 8, causing a lysine-to-alanine exchange at position 232 of the predicted protein (K232A). Alignment of *DGAT1* amino acid sequences of different plant and animal species indicates a conserved lysine residue at position 232 of the bovine sequence (Fig. 2). All other variants are silent, occurring either within introns or in the 5' or 3' untranslated regions.

We unambiguously derived at least eight haplotypes through direct sequencing of animals belonging to different breeds of *B. taurus taurus* and *B. taurus indicus* as well as of animals of *Bos grunniens* (yak) and *Bubalus bubalus* (water buffalo) at six of the variable nucleotide positions. Lysine-encoding haplotypes are present in yak and water buffalo. Thus, the lysine-encoding variant is likely the ancestral state of *DGAT1*. The K232A substitution probably took place early in the history of domesticated cattle or even before domestication as surmised by the presence of the alanine variant in the Anatolian Black breed which is indigenous to a region known as the site of domestication of the European *Bos taurus* (24). The alanine variant was not seen in the subspecies *B. taurus indicus* which was domesticated separately (25). An RFLP assay was applied to obtain preliminary estimates on the frequency of the lysine- and alanine-encoding alleles in several cattle breeds and species of the *Bovinae* subfamily (Fig. 3B).

The allele frequencies at six variable positions in the pools of animals with high and low breeding values for milk fat content,

respectively, are visualized in Fig. 4. Distinct frequency differences exist for the Fleckvieh and Holstein–Friesian breeds between the groups with high and low breeding values. The most extreme differences are between the “positive” and “negative” pools in the Holstein–Friesian breed. In both breeds, the lysine-encoding variant is more frequent in animals with high breeding values. The lysine-encoding allele is also present in Braunvieh animals from the high end of the distribution of the milk fat content breeding values, although to a lesser extent than in the other breeds. This finding is remarkable considering the low frequency of the lysine allele in this breed (Fig. 3B). Allele frequencies calculated for position 10433 by using the genotyping results from the RFLP assay in individual animals of the pools (data not shown) were in good agreement with those estimated from the analysis of the sequencing traces, providing support for the latter.

Association of variation in *DGAT1* with milk fat content variation is evidence that *DGAT1* represents the milk fat QTL on chromosome 14 and that the predicted K232A substitution is responsible for the QTL variation, if the observed association (*i*) is not a relict of population admixture nor (*ii*) is due to a causal variant located in another gene but in linkage disequilibrium with the *DGAT1* variants. Linkage disequilibrium could indeed explain association of *DGAT1* variants and milk fat content, especially if the QTL is confirmed to reside proximal to *ILSTS039* as suggested by Riquet *et al.* (7). Admixture also cannot be ruled out for all three breeds investigated when considering the frequencies of the lysine-encoding allele (Fig. 3B) and the history of the breeds. Admixture in Holstein–Friesian and Braunvieh could result from crossbreeding with Jersey, a breed with high milk fat content and high frequency of the lysine-encoding allele, whereas admixture in Fleckvieh could result from the introgression of Holstein–Friesian. However, pedigree analysis reveals that animals with Holstein–Friesian ancestry can

Table 1. Variable positions in and around *DGAT1*

Position*	Variation	Position*	Variation
1465	(AGGCCCCGCCCTCCCCGG) _n [†]	10515	(G) _n [‡]
3343	C → G/T	≈ 10800	+ → - [§]
3399	T → G	11030	G → A
7233	A → G	11048	C → T
8567	A → G	11993	T → C
8607	G → A	12005	A → C
9284	T → C	12036	T → C
10147	A → C	12056	A → G
10433	G → A	12136	G → A
10434	C → A	13309	G → C

*Numbering is according to the deposited sequence (GenBank accession no. AJ318490).

[†]*n* = 3, 4, 5, 6, 7.

[‡]*n* = 6, 7.

[§]+, PCR product present; -, no or very weak PCR product.

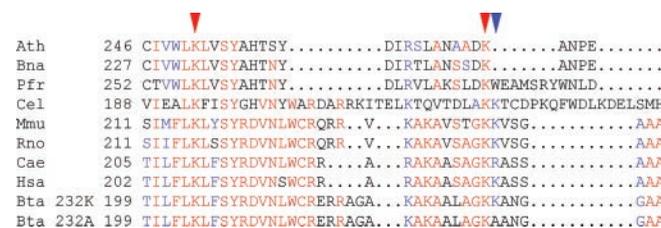


Fig. 2. Alignment of the DGAT amino acid sequences of *Arabidopsis thaliana* (Ath), *Brassica napus* (Bna), *Perilla frutescens* (Pfr), *Caenorhabditis elegans* (Cel), *Mus musculus* (Mmu), *Rattus norvegicus* (Rno), *Ceropithecus aethiops* (Cae), *Homo sapiens* (Hsa) and two alleles of *Bos taurus* (Bta. 232K, Bta.232A) by using PILEUP of the GCG package (23). Sequences were assembled by using BOXSHADE (http://www.isrec.isb-sib.ch:8080/software/BOX_form.html). Numbers on the left indicate amino acid positions. Red and blue letters indicate identical and conserved amino acids, respectively. Red arrows indicate identical lysine residues that might play a role in acyl-CoA binding. The blue arrow indicates conserved amino acids in animal species and in the bovine allele associated with high milk fat content. The lysine-to-alanine mutation at this position is not conservative and could have a negative effect on the acyl-CoA-binding capacity of DGAT.

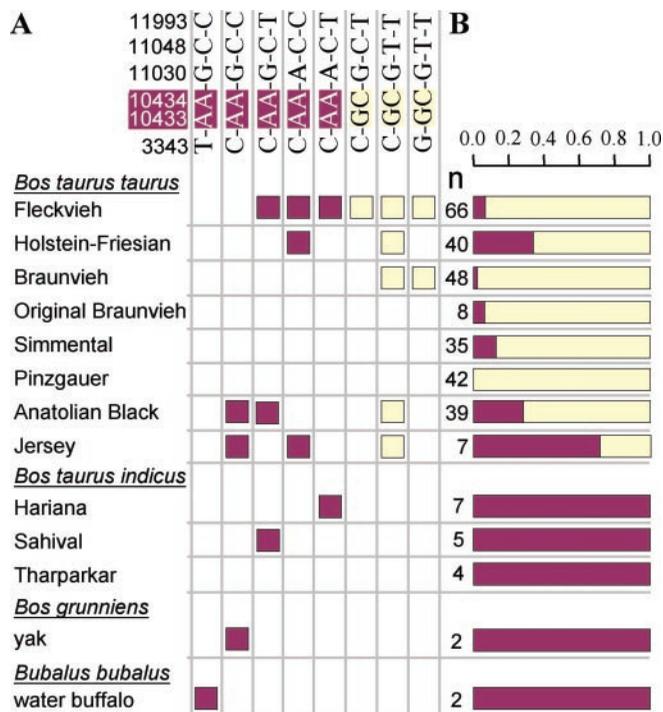


Fig. 3. (A) Haplotypes of *DGAT1* based on nucleotide positions 3343, 10433, 10434, 11030, 11048, and 11993 determined by direct sequencing and (B) preliminary frequency estimates for the lysine (red) and alanine (yellow) encoding alleles in different species and breeds determined by RFLP assay (n, number of animals).

be found both in the “positive” and the “negative” pools, which argues against admixture being exclusively responsible for the observed association.

Another argument for *DGAT1* (or linked loci) being responsible for the QTL variation on chromosome 14 is provided by QTL mapping in Fleckvieh by using a half-sib or “granddaughter” design. The test statistic indicates that the most likely position of the QTL along chromosome 14 is close to marker *ILSTS039* (Fig. 5A). Evidence was highly significant for segregation of the QTL in 2 of 20 families (Fig. 5B) with estimates of QTL effects for milk fat content being 0.313 ± 0.070 and 0.409 ± 0.064 , respectively. These allele substitution effects greatly exceed the genetic standard deviation of about 0.2 in the Fleckvieh population and account for about 10% of the phenotypic variation in this breed (based on a frequency of 0.07). The genotypes at the predicted K232A substitution determined by an RFLP assay are compatible with the heterozygous status of the segregating (*Qq*) sires and homozygosity of the alanine-encoding variant of the nonsegregating (most likely *qq*) sires (Fig. 5C). Direct sequencing of *DGAT1* from DNA and determining the repeat number of the 5' VNTR in the two segregating bulls and some of their progeny allowed us to derive their haplotypes on the basis of the genotypes of homozygous progeny. The lysine-encoding variant is present on two different haplotypes. Bull 16 carries the only lysine-encoding haplotype we found in Holstein-Friesian, and bull 10 carries a lysine-encoding haplotype found in Fleckvieh, Anatolian Black, and Sahival (Figs. 3A and 6), which could indicate that a lysine-encoding allele has been introduced into Fleckvieh from Holstein-Friesian. Pedigree analysis indeed shows that the greatgrandfather of bull 16 was a purebred Holstein-Friesian sire, whereas no indication of Hol-

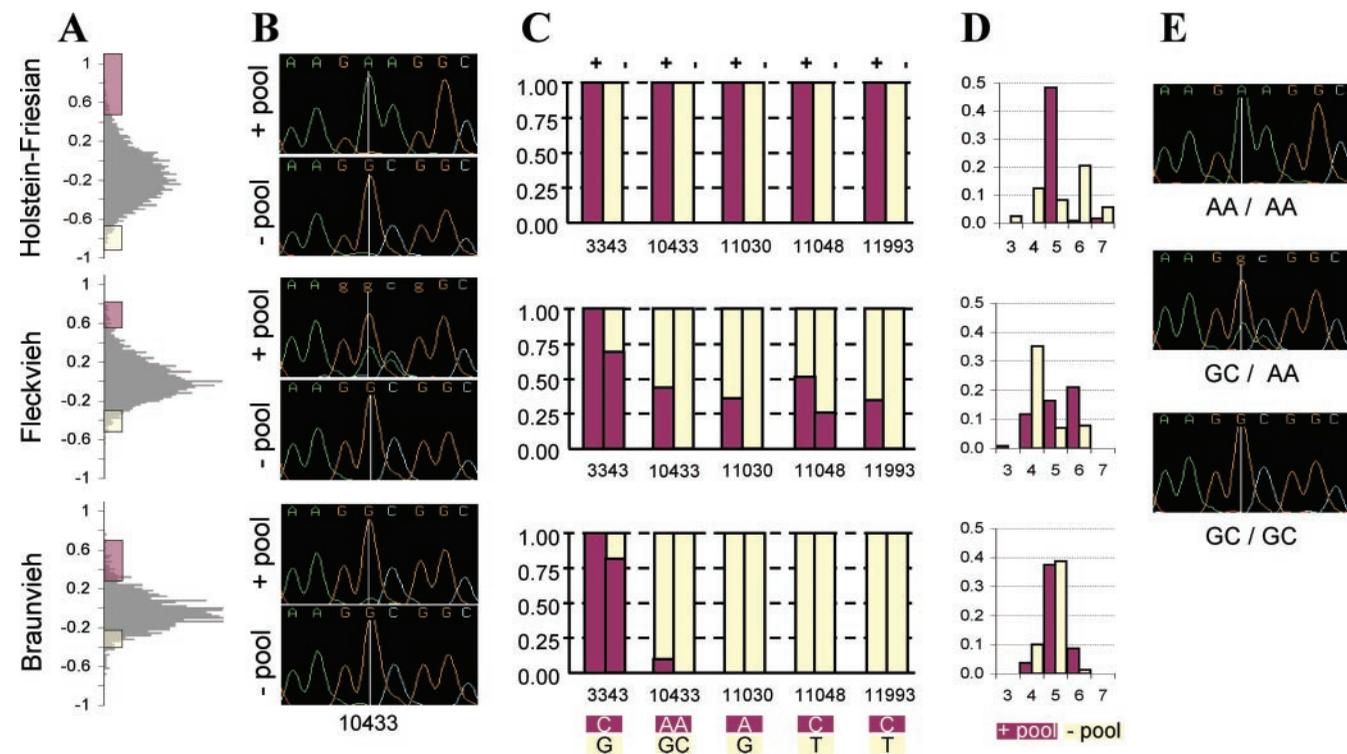


Fig. 4. (A) Distributions of breeding values for milk fat content of Holstein-Friesian, Fleckvieh, and Braunvieh artificial insemination bulls. Colored areas indicate the range of the breeding values, from which bulls were chosen for the extreme positive (+, red) and negative (–, yellow) pools. (B and E) Sequence trace views of sequencing traces for positions 10430–10437 within *DGAT1* for DNA pools (B) and individual animals (E). The vertical line indicates nucleotide position 10433. Positions 10433 and 10434 are responsible for the K232A substitution. (C) Estimates of allelic frequencies in the +pool (red) and –pool (yellow) for each breed based on sequencing traces. (D) Frequencies of alleles with 3, 4, 5, 6, and 7 repeat units in the 5' region of *DGAT1* in the +pool (red) and –pool (yellow) for each breed determined by fragment analysis.

represent the more efficient allele with regard to fat synthesis, because it is present at a high frequency or exclusively in species and breeds with a higher milk fat content and in animals from the “high-end” of the trait distribution.

Proving the identity of the QTL alleles with *DGATI* alleles by genetic means, i.e., by exclusion of linkage disequilibrium, which is *a priori* extensive in cattle (28), will be difficult. At the very least (29), we have revised the map position of the QTL reported by Riquet *et al.* (7) during the final stage of preparing this manuscript to encompass a region distal to *ILSTS039*. Thus, the map position of *DGATI* no longer precludes it theoretically from containing causal QTL variation. However, proving a causal involvement of *DGATI* will require studying the effects of the different allelic variants on the enzymatic activity both *in vitro* and *in vivo*, i.e., in transgenic animals. Our genetic studies provide a basis to establish hypotheses concerning determinants affecting the rate of fat synthesis not only in the context of triglyceride production in animals and plants but also in the context of human diseases caused by excess triglyceride synthesis and deposition.

If the K232A polymorphism is indeed identical with allelic variation at the milk fat QTL, it represents “classical” QTL variation, i.e., the molecular variation of *DGATI* would not be the only explanation for genetic variation of milk fat content. Such variation is compatible with the model of quantitative genetics assuming variation is due to so-called “polygenes” responsible for a continuously varying trait. The identification of causal molecular variation in one of the polygenes should facilitate the molecular analysis of other genes involved in milk fat content variation.

In our QTL study we assumed a biallelic situation, i.e., the alleles *Q* and *q*. The presence of both the milk fat content-enhancing and -lowering variants in different haplotypes indicates the presence of multiple alleles at the QTL. This hypothesis can now be tested by analyzing the effects of the different haplotypes on milk fat content. It will be particularly interesting to involve the VNTR in the 5′ region in such studies. Although variation in the number of repeats is unlikely to be a major causal determinant, especially considering that the five- and six-repeat alleles occur equally frequently in the two Braunvieh pools (Fig. 4D), it remains that the VNTR could affect the regulation of the gene. Together, variation in the expression level of the gene and in the activity of the enzyme due to a polymorphism affecting the amino acid sequence could determine the net rate of triglyceride synthesis.

We thank the Bayerische Landesanstalt für Tierzucht in Grub and the Vereinigte Informationssysteme für Tierhaltung w.V. in Verden for providing pedigree information and INTERBULL breeding values. Semen samples were generously donated by German artificial insemination stations. C. Kühn assembled semen samples for Holstein–Friesian pools. DNA used in pedigree analyses was prepared by K. Bodis. T. Meitinger supplied DNA samples of radiation hybrid panel GB4. R. Antes, R. Böhm, and A. Keller provided excellent technical assistance. H. Gossner contributed valuable comments concerning the Fleckvieh pedigrees. Zebu DNA was generously provided by D. Bradley. O. Bininda-Emonds provided valuable comments on the manuscript. This work was funded by the Arbeitsgemeinschaft Deutscher Rinderzüchter (German cattle breeders federation), the Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research, Project 0311020A), and the European Union (EURIBDIS Project, BIO 4-CT97-0471).

- Goddard, M. E. & Wiggans, G. R. (1999) in *The Genetics of Cattle*, eds. Fries, R. & Ruvinsky, A. (CABI, Wallingford, U.K.), pp. 511–537.
- Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A. T., Sargent, L. S., Sorensen, A., Steele, M. R., Zahor, X., *et al.* (1995) *Genetics* **139**, 907–920.
- Zhang, Q., Boichard, D., Hoeschele, I., Ernst, C., Eggen, A., Murkve, B., Pfister-Genskow, M., Witte, L. A., Grignola, F. E., Uimari, P., *et al.* (1998) *Genetics* **149**, 1959–1973.
- Heyen, D. W., Weller, J. I., Ron, M., Band, M., Beever, J. E., Feldmesser, E., Da, Y., Wiggans, G. R., VanRanden, P. M. & Lewin, H. A. (1999) *Physiol. Genomics* **1**, 165–175.
- Velmalu, R. J., Vilkkki, H. J., Elo, K. T., de Koning, D. J. & Maki-Tanila, A. V. (1999) *Anim. Genet.* **30**, 136–143.
- Coppieters, W., Riquet, J., Arranz, J. J., Berzi, P., Cambisano, N., Grisart, B., Karim, L., Marcq, F., Moreau, L., Nezer, C., *et al.* (1998) *Mamm. Genome* **9**, 540–544.
- Riquet, J., Coppieters, W., Cambisano, N., Arranz, J. J., Berzi, P., Davis, S. K., Grisart, B., Farnir, F., Karim, L., Mni, M., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96**, 9252–9257.
- Cases, S., Stone, S. J., Zhou, P., Yen, E., Tow, B., Lardizabal, K. D., Voelker, T. & Farese, R. V., Jr. (2001) *J. Biol. Chem.* **276**, 38870–38876.
- Smith, S. J., Cases, S., Jensen, D. R., Chen, H. C., Sande, E., Tow, B., Sanan, D. A., Raber, J., Eckel, R. H. & Farese, R. V., Jr. (2000) *Nat. Genet.* **25**, 87–90.
- Warren, W., Smith, T. P., Rexroad, C. E., Fahrenkrug, S. C., Allison, T., Shu, C. L., Catanese, J. & de Jong, P. J. (2000) *Mamm. Genome* **11**, 662–663.
- Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vegaczary, N., Spillet, D., Mueslet, D., Prud’Homme, J. F., Dib, C., Auffray, C., *et al.* (1996) *Hum. Mol. Genet.* **5**, 339–346.
- Womack, J. E., Johnson, J. S., Owens, E. K., Rexroad, C. E., III, Schläpfer, J. & Yang, Y.-P. (1997) *Mamm. Genome* **8**, 854–856.
- Henderson, C. R. (1974) *J. Dairy Sci.* **57**, 963–972.
- Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) *Genome Res.* **8**, 175–185.
- Ewing, B. & Green, P. (1998) *Genome Res.* **8**, 186–194.
- Nickerson, D. A., Tobe, V. O. & Taylor, S. L. (1997) *Nucleic Acids Res.* **25**, 2745–2751.
- Gordon, D., Abajian, C. & Green, P. (1998) *Genome Res.* **8**, 195–202.
- Weller, J. I., Kashi, Y. & Soller, M. (1990) *J. Dairy Sci.* **73**, 2525–2537.
- Knott, S. A., Elsen, J. M. & Haley, C. S. (1996) *Theor. Appl. Genet.* **93**, 71–80.
- Churchill, G. A. & Doerge, R. W. (1994) *Genetics* **138**, 963–971.
- Snedecor, G. W. & Cochran, W. G. (1980) *Statistical Methods* (Iowa State Univ. Press, Ames).
- Rebeiz, M. & Lewin, H. A. (2000) *Anim. Biotechnol.* **11**, 75–241.
- Genetics Computer Group (2001) WISCONSIN PACKAGE (GCG, Madison, WI), Version 10.2.
- Medjugorac, I., Kusterman, W., Lazar, P., Russ, I. & Pirchner, F. (1994) *Anim. Genet.* **25**, 19–27.
- Loftus, R. T., Machugh, D. E., Bradley, D. G., Sharp, P. M. & Cunningham, P. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2757–2761.
- Mathieu, M., Modis, Y., Zeelen, J. P., Engel, C. K., Abagyan, R. A., Ahlberg, A., Rasmussen, B., Lamzin, V. S., Kunau, W. H. & Wierenga, R. K. (1997) *J. Mol. Biol.* **273**, 714–728.
- Mukherjee, S. & Katiyar, S. S. (2000) *J. Enzyme Inhib.* **15**, 421–427.
- Farnir, F., Coppieters, W., Arranz, J.-J., Berzi, P., Cambisano, N., Grisart, B., Karim, L., Marcq, F., Moreau, L., Mni, M., *et al.* (2000) *Genome Res.* **10**, 220–227.
- Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., *et al.* (2002) *Genome Res.* **12**, 222–231.